Appendix I: Alignment of instant SEQ ID NO: 1 and Q9H4I2, publicly available June 2002

BLASTP 2.2.22+
Reference: Stephen F. Altschul, Thomas L. Madden, Alejandro
A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and
David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new
generation of protein database search programs", Nucleic
Acids Res. 25:3389-3402.

Reference for compositional score matrix adjustment: Stephen
F. Altschul, John C. Wootton, E. Michael Gertz, Richa
Agarwala, Aleksandr Morgulis, Alejandro A. Schaffer, and
Yi-Kuo Yu (2005) "Protein database searches using
compositionally adjusted substitution matrices", FEBS J.
272:5101-5109.

RID: KMB3KFZ911R

Query= SID_1
Length=956


```
                                                        Score      E
Sequences producing significant alignments:            (Bits)   Value

1cl|22739  TrEMBL|Q9H4I2|Release 21|01-JUN-2002          1989    0.0

ALIGNMENTS
>1cl|22739 TrEMBL|Q9H4I2|Release 21|01-JUN-2002
Length=956

 Score = 1989 bits (5152),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 956/956 (100%), Positives = 956/956 (100%), Gaps = 0/956 (0%)

Query  1    MASKRKSTTPCMIPVKTVVLQDASMEAQPAETLPEGPQQDLPPEASAASSEAAQNPSSTD  60
            MASKRKSTTPCMIPVKTVVLQDASMEAQPAETLPEGPQQDLPPEASAASSEAAQNPSSTD
Sbjct  1    MASKRKSTTPCMIPVKTVVLQDASMEAQPAETLPEGPQQDLPPEASAASSEAAQNPSSTD  60

Query  61   GSTLANGHRSTLDGYLYSCKYCDFRSHDMTQFVGHMNSEHTDFNKDPTFVCSGCSFLAKT  120
            GSTLANGHRSTLDGYLYSCKYCDFRSHDMTQFVGHMNSEHTDFNKDPTFVCSGCSFLAKT
Sbjct  61   GSTLANGHRSTLDGYLYSCKYCDFRSHDMTQFVGHMNSEHTDFNKDPTFVCSGCSFLAKT  120

Query  121  PEGLSLHNATCHSGEASFVWNVAKPDNHVVVEQSIPESTSTPDLAGEPSAEGADGQAEII  180
            PEGLSLHNATCHSGEASFVWNVAKPDNHVVVEQSIPESTSTPDLAGEPSAEGADGQAEII
Sbjct  121  PEGLSLHNATCHSGEASFVWNVAKPDNHVVVEQSIPESTSTPDLAGEPSAEGADGQAEII  180

Query  181  ITKTPIMKIMKGKAEAKKIHTLKENVPSQPVGEALPKLSTGEMEVREGDHSFINGAVPVS  240
            ITKTPIMKIMKGKAEAKKIHTLKENVPSQPVGEALPKLSTGEMEVREGDHSFINGAVPVS
Sbjct  181  ITKTPIMKIMKGKAEAKKIHTLKENVPSQPVGEALPKLSTGEMEVREGDHSFINGAVPVS  240

Query  241  QASASSAKNPHAANGPLIGTVPVLPAGIAQFLSLQQQPPVHAQHHVHQPLPTAKALPKVM  300
            QASASSAKNPHAANGPLIGTVPVLPAGIAQFLSLQQQPPVHAQHHVHQPLPTAKALPKVM
Sbjct  241  QASASSAKNPHAANGPLIGTVPVLPAGIAQFLSLQQQPPVHAQHHVHQPLPTAKALPKVM  300

Query  301  IPLSSIPTYNAAMDSNSFLKNSFHKFPYPTKAELCYLTVVTKYPEEQLKIWFTAQRLKQG  360
            IPLSSIPTYNAAMDSNSFLKNSFHKFPYPTKAELCYLTVVTKYPEEQLKIWFTAQRLKQG
Sbjct  301  IPLSSIPTYNAAMDSNSFLKNSFHKFPYPTKAELCYLTVVTKYPEEQLKIWFTAQRLKQG  360

Query  361  ISWSPEEIEDARKKMFNTVIQSVPQPTITVLNTPLVASAGNVQHLIQAALPGHVVGQPEG  420
```

```
                 ISWSPEEIEDARKKMFNTVIQSVPQPTITVLNTPLVASAGNVQHLIQAALPGHVVGQPEG
Sbjct   361      ISWSPEEIEDARKKMFNTVIQSVPQPTITVLNTPLVASAGNVQHLIQAALPGHVVGQPEG      420

Query   421      TGGGLLVTQPLMANGLQATSSPLPLTVTSVPKQPGVAPINTVCSNTTSAVKVVNAAQSLL      480
                 TGGGLLVTQPLMANGLQATSSPLPLTVTSVPKQPGVAPINTVCSNTTSAVKVVNAAQSLL
Sbjct   421      TGGGLLVTQPLMANGLQATSSPLPLTVTSVPKQPGVAPINTVCSNTTSAVKVVNAAQSLL      480

Query   481      TACPSITSQAFLDASIYKNKKSHEQLSALKGSFCRNQFPGQSEVEHLTKVTGLSTREVRK      540
                 TACPSITSQAFLDASIYKNKKSHEQLSALKGSFCRNQFPGQSEVEHLTKVTGLSTREVRK
Sbjct   481      TACPSITSQAFLDASIYKNKKSHEQLSALKGSFCRNQFPGQSEVEHLTKVTGLSTREVRK      540

Query   541      WFSDRRYHCRNLKGSRAMIPGDHSSIIIDSVPEVSFSPSSKVPEVTCIPTTATLATHPSA      600
                 WFSDRRYHCRNLKGSRAMIPGDHSSIIIDSVPEVSFSPSSKVPEVTCIPTTATLATHPSA
Sbjct   541      WFSDRRYHCRNLKGSRAMIPGDHSSIIIDSVPEVSFSPSSKVPEVTCIPTTATLATHPSA      600

Query   601      KRQSWHQTPDFTPTKYKERAPEQLRALESSFAQNPLPLDEELDRLRSETKMTRREIDSWF      660
                 KRQSWHQTPDFTPTKYKERAPEQLRALESSFAQNPLPLDEELDRLRSETKMTRREIDSWF
Sbjct   601      KRQSWHQTPDFTPTKYKERAPEQLRALESSFAQNPLPLDEELDRLRSETKMTRREIDSWF      660

Query   661      SERRKKVNAEETKKAEENASQEEEEAAEDEGGEEDLASELRVSGENGSLEMPSSHILAER      720
                 SERRKKVNAEETKKAEENASQEEEEAAEDEGGEEDLASELRVSGENGSLEMPSSHILAER
Sbjct   661      SERRKKVNAEETKKAEENASQEEEEAAEDEGGEEDLASELRVSGENGSLEMPSSHILAER      720

Query   721      KVSPIKINLKNLRVTEANGRNEIPGLGACDPEDDESNKLAEQLPGKVSCKKTAQQRHLLR      780
                 KVSPIKINLKNLRVTEANGRNEIPGLGACDPEDDESNKLAEQLPGKVSCKKTAQQRHLLR
Sbjct   721      KVSPIKINLKNLRVTEANGRNEIPGLGACDPEDDESNKLAEQLPGKVSCKKTAQQRHLLR      780

Query   781      QLFVQTQWPSNQDYDSIMAQTGLPRPEVVRWFGDSRYALKNGQLKWYEDYKRGNFPPGLL      840
                 QLFVQTQWPSNQDYDSIMAQTGLPRPEVVRWFGDSRYALKNGQLKWYEDYKRGNFPPGLL
Sbjct   781      QLFVQTQWPSNQDYDSIMAQTGLPRPEVVRWFGDSRYALKNGQLKWYEDYKRGNFPPGLL      840

Query   841      VIAPGNRELLQDYYMTHKMLYEEDLQNLCDKTQMSSQQVKQWFAEKMGEETRAVADTGSE      900
                 VIAPGNRELLQDYYMTHKMLYEEDLQNLCDKTQMSSQQVKQWFAEKMGEETRAVADTGSE
Sbjct   841      VIAPGNRELLQDYYMTHKMLYEEDLQNLCDKTQMSSQQVKQWFAEKMGEETRAVADTGSE      900

Query   901      DQGPGTGELTAVHKGMGDTYSEVSENSESWEPRVPEASSEPFDTSSPQAGRQLETD       956
                 DQGPGTGELTAVHKGMGDTYSEVSENSESWEPRVPEASSEPFDTSSPQAGRQLETD
Sbjct   901      DQGPGTGELTAVHKGMGDTYSEVSENSESWEPRVPEASSEPFDTSSPQAGRQLETD       956
```